

1636



RAW SEQUENCE LISTING

DATE: 08/20/2002

PATENT APPLICATION: US/09/856,319B

TIME: 14:39:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\08202002\I856319B.raw

3 <110> APPLICANT: UEMURA, Hidetoshi
 4 OKUI, Akira
 5 KOMINAMI, Katsuya
 6 YAMAGUCHI, Nozomi
 7 MITSUI, Shinichi
 9 <120> TITLE OF INVENTION: NOVEL SERINE PROTEASE BSSP5
 11 <130> FILE REFERENCE: UEMURA=5
 13 <140> CURRENT APPLICATION NUMBER: 09/856,319B
 14 <141> CURRENT FILING DATE: 2001-05-21
 16 <150> PRIOR APPLICATION NUMBER: JP 10/347806
 17 <151> PRIOR FILING DATE: 1998-11-20
 19 <150> PRIOR APPLICATION NUMBER: PCT JP99/06473
 20 <151> PRIOR FILING DATE: 1999-11-19
 22 <160> NUMBER OF SEQ ID NOS: 32
 24 <170> SOFTWARE: PatentIn version 3.1
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 1149
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Homo sapiens
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (11)..(802)
 34 <223> OTHER INFORMATION:
 36 <220> FEATURE:
 37 <221> NAME/KEY: mat_peptide
 38 <222> LOCATION: (110)..()
 39 <223> OTHER INFORMATION:
 41 <400> SEQUENCE: 1
 42 atctgccacg atg ttg ctg ctc agc ctg acc cta agc ctg gtt ctc ctc 49
 43 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu
 44 -30 -25
 46 ggc tcc tcc tgg ggc tgc ggc att cct gcc atc aaa ccg gca ctg agc 97
 47 Gly Ser Ser Trp Gly Cys Gly Ile Pro Ala Ile Lys Pro Ala Leu Ser
 48 -20 -15 -10 -5
 50 ttc agc cag agg att gtc aac ggg gag aat gca gtg ttg ggc tcc tgg 145
 51 Phe Ser Gln Arg Ile Val Asn Gly Glu Asn Ala Val Leu Gly Ser Trp
 52 -1 1 5 10
 54 ccc tgg cag gtg tcc ctg cag gac agc agc ggc ttc cac ttc tgc ggt 193
 55 Pro Trp Gln Val Ser Leu Gln Asp Ser Ser Gly Phe His Phe Cys Gly
 56 15 20 25
 58 ggt tct ctc atc agc cag tcc tgg gtg gtc act gct gcc cac tgc aat 241
 59 Gly Ser Leu Ile Ser Gln Ser Trp Val Val Thr Ala Ala His Cys Asn
 60 30 35 40

ENTERED

RAW SEQUENCE LISTING

DATE: 08/20/2002

PATENT APPLICATION: US/09/856,319B

TIME: 14:39:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\08202002\I856319B.raw

```

62 gtc agc cct ggc cgc cat ttt gtt gtc ctg ggc gag tat gac cga tca      289
63 Val Ser Pro Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser
64 45                               50                               55                               60
66 tca aac gca gag ccc ttg cag gtt ctg tcc gtc tct cgg gcc att aca      337
67 Ser Asn Ala Glu Pro Leu Gln Val Leu Ser Val Ser Arg Ala Ile Thr
68                               65                               70                               75
70 cac cct agc tgg aac tct acc acc atg aac aat gac gtg acg ctg ctg      385
71 His Pro Ser Trp Asn Ser Thr Thr Met Asn Asn Asp Val Thr Leu Leu
72                               80                               85                               90
74 aag ctc gcc tcg cca gcc cag tac aca aca cgc atc tcg cca gtt tgc      433
75 Lys Leu Ala Ser Pro Ala Gln Tyr Thr Thr Arg Ile Ser Pro Val Cys
76                               95                               100                              105
78 ctg gca tcc tca aac gag gct ctg act gaa ggc ctc acg tgt gtc acc      481
79 Leu Ala Ser Ser Asn Glu Ala Leu Thr Glu Gly Leu Thr Cys Val Thr
80                               110                               115                               120
82 acc ggc tgg ggt cgc ctc agt ggc gtg ggc aat gtg aca cca gca cat      529
83 Thr Gly Trp Gly Arg Leu Ser Gly Val Gly Asn Val Thr Pro Ala His
84 125                               130                               135                               140
86 ctg cag cag gtg gct ttg ccc ctg gtc act gtg aat cag tgc cgg cag      577
87 Leu Gln Gln Val Ala Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln
88                               145                               150                               155
90 tac tgg gac tca agt atc act gac tcc atg atc tgt gca ggt ggc gca      625
91 Tyr Trp Asp Ser Ser Ile Thr Asp Ser Met Ile Cys Ala Gly Gly Ala
92                               160                               165                               170
94 ggt gcc tcc tcg tgc cag ggt gac tcc gga ggc cct ctt gtc tgc cag      673
95 Gly Ala Ser Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln
96                               175                               180                               185
98 aag gga aac aca tgg gtg ctt att ggt att gtc tcc tgg ggc acc aaa      721
99 Lys Gly Asn Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys
100                               190                               195                               200
102 aac tgc aat gtg cgc gca cct gct gtg tat act cga gtt agc aag ttc      769
103 Asn Cys Asn Val Arg Ala Pro Ala Val Tyr Thr Arg Val Ser Lys Phe
104 205                               210                               215                               220
106 agc acc tgg atc aac cag gtc ata gcc tac aac tgagctcacc acagggcctc      822
107 Ser Thr Trp Ile Asn Gln Val Ile Ala Tyr Asn
108                               225                               230
110 cccagctcaa cccatttaaa ggacccaggc cctgtcccat catgcattca tgtctgtctt      882
112 cctggctcag gagaaagaag aggctgttga gggctccgact ccctacttgg acttctggca      942
114 cagaaggggc tgagtgaact cttgagtagc agtggctctt cctagagtag ccatgccgtg      1002
116 gccggggccc ccacccctcc tccagggcaa ccccttggtc ctacagcaag aagccagaac      1062
118 tgttggaatg aatggcagcc ctccttgag aggcagcctg tttactgaat acagaggata      1122
120 cgtttacaaa aaaaaaaaaa aaaaaaaa      1149
123 <210> SEQ ID NO: 2
124 <211> LENGTH: 264
125 <212> TYPE: PRT
126 <213> ORGANISM: Homo sapiens
128 <400> SEQUENCE: 2
130 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu Gly Ser Ser
131                               -30                               -25                               -20

```

RAW SEQUENCE LISTING

DATE: 08/20/2002

PATENT APPLICATION: US/09/856,319B

TIME: 14:39:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\08202002\I856319B.raw

```

134 Trp Gly Cys Gly Ile Pro Ala Ile Lys Pro Ala Leu Ser Phe Ser Gln
135          -15          -10          -5
138 Arg Ile Val Asn Gly Glu Asn Ala Val Leu Gly Ser Trp Pro Trp Gln
139 -1 1          5          10          15
142 Val Ser Leu Gln Asp Ser Ser Gly Phe His Phe Cys Gly Gly Ser Leu
143          20          25          30
146 Ile Ser Gln Ser Trp Val Val Thr Ala Ala His Cys Asn Val Ser Pro
147          35          40          45
150 Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser Ser Asn Ala
151          50          55          60
154 Glu Pro Leu Gln Val Leu Ser Val Ser Arg Ala Ile Thr His Pro Ser
155          65          70          75
158 Trp Asn Ser Thr Thr Met Asn Asn Asp Val Thr Leu Leu Lys Leu Ala
159 80          85          90          95
162 Ser Pro Ala Gln Tyr Thr Thr Arg Ile Ser Pro Val Cys Leu Ala Ser
163          100          105          110
166 Ser Asn Glu Ala Leu Thr Glu Gly Leu Thr Cys Val Thr Thr Gly Trp
167          115          120          125
170 Gly Arg Leu Ser Gly Val Gly Asn Val Thr Pro Ala His Leu Gln Gln
171          130          135          140
174 Val Ala Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln Tyr Trp Asp
175          145          150          155
178 Ser Ser Ile Thr Asp Ser Met Ile Cys Ala Gly Gly Ala Gly Ala Ser
179 160          165          170          175
182 Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Lys Gly Asn
183          180          185          190
186 Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys Asn Cys Asn
187          195          200          205
190 Val Arg Ala Pro Ala Val Tyr Thr Arg Val Ser Lys Phe Ser Thr Trp
191          210          215          220
194 Ile Asn Gln Val Ile Ala Tyr Asn
195          225          230
198 <210> SEQ ID NO: 3
199 <211> LENGTH: 834
200 <212> TYPE: DNA
201 <213> ORGANISM: Mus sp.
203 <220> FEATURE:
204 <221> NAME/KEY: mat_peptide
205 <222> LOCATION: (132)..()
206 <223> OTHER INFORMATION:
208 <220> FEATURE:
209 <221> NAME/KEY: CDS
210 <222> LOCATION: (33)..(824)
211 <223> OTHER INFORMATION:
213 <400> SEQUENCE: 3
214 gaccatctca acaccattcc ttatttgtca ca atg cta ctg ctc agc cta acc      53
215                               Met Leu Leu Leu Ser Leu Thr
216                               -30
218 ctt agc ctg gtc ctc ctt ggc tcc tcc tgg ggc tgt ggt gtt cct gcc      101

```

RAW SEQUENCE LISTING

DATE: 08/20/2002

PATENT APPLICATION: US/09/856,319B

TIME: 14:39:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\08202002\I856319B.raw

219	Leu	Ser	Leu	Val	Leu	Leu	Gly	Ser	Ser	Trp	Gly	Cys	Gly	Val	Pro	Ala	
220		-25					-20				-15						
222	atc	acg	cct	gca	ctg	agc	tac	aat	cag	aga	att	gtc	aac	ggg	gag	aat	149
223	Ile	Thr	Pro	Ala	Leu	Ser	Tyr	Asn	Gln	Arg	Ile	Val	Asn	Gly	Glu	Asn	
224	-10					-5			-1	1				5			
226	gca	gtg	cca	ggc	tcc	tgg	ccc	tgg	cag	gtg	tct	ctc	cag	gat	aac	acc	197
227	Ala	Val	Pro	Gly	Ser	Trp	Pro	Trp	Gln	Val	Ser	Leu	Gln	Asp	Asn	Thr	
228				10					15				20				
230	ggc	ttc	cac	ttc	tgc	ggt	ggt	tct	ctc	atc	agt	ccg	aac	tgg	gtg	gtc	245
231	Gly	Phe	His	Phe	Cys	Gly	Gly	Ser	Leu	Ile	Ser	Pro	Asn	Trp	Val	Val	
232			25					30					35				
234	acg	gct	gcc	cac	tgc	caa	gtc	acg	cct	gga	cgc	cac	ttt	gtc	gtt	ttg	293
235	Thr	Ala	Ala	His	Cys	Gln	Val	Thr	Pro	Gly	Arg	His	Phe	Val	Val	Leu	
236		40					45					50					
238	gga	gaa	tat	gac	cga	tct	tcc	aat	gct	gaa	cct	gtg	cag	gtc	ctc	tcg	341
239	Gly	Glu	Tyr	Asp	Arg	Ser	Ser	Asn	Ala	Glu	Pro	Val	Gln	Val	Leu	Ser	
240	55					60					65				70		
242	atc	gca	agg	gcc	atc	aca	cac	cct	aac	tgg	aac	gcc	aac	acc	atg	aac	389
243	Ile	Ala	Arg	Ala	Ile	Thr	His	Pro	Asn	Trp	Asn	Ala	Asn	Thr	Met	Asn	
244				75						80					85		
246	aat	gac	ctg	act	ctc	ctg	aag	ctt	gcc	tcg	cca	gcc	cgg	tac	aca	gca	437
247	Asn	Asp	Leu	Thr	Leu	Leu	Lys	Leu	Ala	Ser	Pro	Ala	Arg	Tyr	Thr	Ala	
248			90						95					100			
250	caa	gtc	tca	cca	gtc	tgc	ctg	gct	tcc	aca	aac	gag	gca	ctg	cct	tcg	485
251	Gln	Val	Ser	Pro	Val	Cys	Leu	Ala	Ser	Thr	Asn	Glu	Ala	Leu	Pro	Ser	
252			105						110				115				
254	ggg	ctc	acc	tgt	gtc	acc	act	ggc	tgg	ggc	cga	atc	agt	ggt	gtg	ggc	533
255	Gly	Leu	Thr	Cys	Val	Thr	Thr	Gly	Trp	Gly	Arg	Ile	Ser	Gly	Val	Gly	
256		120						125					130				
258	aat	gtg	aca	cca	gct	cgc	ctg	cag	caa	gtt	gtt	cta	ccc	ctg	gtc	act	581
259	Asn	Val	Thr	Pro	Ala	Arg	Leu	Gln	Gln	Val	Val	Leu	Pro	Leu	Val	Thr	
260	135					140						145				150	
262	gtg	aat	cag	tgt	cgg	cag	tac	tgg	ggt	gca	cgc	att	acc	gat	gcc	atg	629
263	Val	Asn	Gln	Cys	Arg	Gln	Tyr	Trp	Gly	Ala	Arg	Ile	Thr	Asp	Ala	Met	
264				155						160					165		
266	ata	tgt	gca	ggt	ggc	tca	ggc	gcc	tcc	tca	tgt	cag	ggt	gac	tca	gga	677
267	Ile	Cys	Ala	Gly	Gly	Ser	Gly	Ala	Ser	Ser	Cys	Gln	Gly	Asp	Ser	Gly	
268			170						175					180			
270	ggc	cct	ctt	gtc	tgc	cag	aag	gga	aac	acc	tgg	gtg	ctt	att	ggg	att	725
271	Gly	Pro	Leu	Val	Cys	Gln	Lys	Gly	Asn	Thr	Trp	Val	Leu	Ile	Gly	Ile	
272			185						190					195			
274	gtc	tcc	tgg	ggc	act	aag	aac	tgc	aac	ata	caa	gca	ccg	gcc	atg	tac	773
275	Val	Ser	Trp	Gly	Thr	Lys	Asn	Cys	Asn	Ile	Gln	Ala	Pro	Ala	Met	Tyr	
276		200						205					210				
278	act	cgg	gtc	agc	aag	ttc	agt	acc	tgg	atc	aac	caa	gtc	atg	gcc	tac	821
279	Thr	Arg	Val	Ser	Lys	Phe	Ser	Thr	Trp	Ile	Asn	Gln	Val	Met	Ala	Tyr	
280	215					220					225				230		
282	aac	taaa	ctgtcc														834
283	Asn																

RAW SEQUENCE LISTING

DATE: 08/20/2002

PATENT APPLICATION: US/09/856,319B

TIME: 14:39:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\08202002\I856319B.raw

```

287 <210> SEQ ID NO: 4
288 <211> LENGTH: 264
289 <212> TYPE: PRT
290 <213> ORGANISM: Mus sp.
292 <400> SEQUENCE: 4
294 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu Gly Ser Ser
295           -30           -25           -20
298 Trp Gly Cys Gly Val Pro Ala Ile Thr Pro Ala Leu Ser Tyr Asn Gln
299           -15           -10           -5
302 Arg Ile Val Asn Gly Glu Asn Ala Val Pro Gly Ser Trp Pro Trp Gln
303 -1 1           5           10           15
306 Val Ser Leu Gln Asp Asn Thr Gly Phe His Phe Cys Gly Gly Ser Leu
307           20           25           30
310 Ile Ser Pro Asn Trp Val Val Thr Ala Ala His Cys Gln Val Thr Pro
311           35           40           45
314 Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser Ser Asn Ala
315           50           55           60
318 Glu Pro Val Gln Val Leu Ser Ile Ala Arg Ala Ile Thr His Pro Asn
319           65           70           75
322 Trp Asn Ala Asn Thr Met Asn Asn Asp Leu Thr Leu Leu Lys Leu Ala
323 80           85           90           95
326 Ser Pro Ala Arg Tyr Thr Ala Gln Val Ser Pro Val Cys Leu Ala Ser
327           100          105          110
330 Thr Asn Glu Ala Leu Pro Ser Gly Leu Thr Cys Val Thr Thr Gly Trp
331           115          120          125
334 Gly Arg Ile Ser Gly Val Gly Asn Val Thr Pro Ala Arg Leu Gln Gln
335           130          135          140
338 Val Val Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln Tyr Trp Gly
339           145          150          155
342 Ala Arg Ile Thr Asp Ala Met Ile Cys Ala Gly Gly Ser Gly Ala Ser
343 160          165          170          175
346 Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Lys Gly Asn
347           180          185          190
350 Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys Asn Cys Asn
351           195          200          205
354 Ile Gln Ala Pro Ala Met Tyr Thr Arg Val Ser Lys Phe Ser Thr Trp
355           210          215          220
358 Ile Asn Gln Val Met Ala Tyr Asn
359           225          230
362 <210> SEQ ID NO: 5
363 <211> LENGTH: 99
364 <212> TYPE: DNA
365 <213> ORGANISM: Artificial Sequence
367 <220> FEATURE:
368 <223> OTHER INFORMATION: Designed oligonucleotide to construct plasmid pSecTrypHis.
370 <400> SEQUENCE: 5
371 aagcttggtgct agcaacacca tgaatctact cctgatcctt. acctttgttg ctgctgctgt      60
373 tgctgcccc tttgacgacg atgacaagga tccgaattc      99
376 <210> SEQ ID NO: 6

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 08/20/2002
PATENT APPLICATION: US/09/856,319B TIME: 14:39:46

Input Set : A:\sequence listing.txt
Output Set: N:\CRF4\08202002\I856319B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:14; N Pos. 9,12
Seq#:15; N Pos. 12,15